example, it has been shown in yeast that genes with similar functions have similar gene expression profiles (Eisen *et al.* (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95, 14863-14868).

The present invention advances the art by providing the predicted transcript sequences (SEQ ID NOS:1-39010), for 39010 transcripts predicted from the assembled human genome, many of which did not have evidence for their existence in the prior art. Furthermore, the present invention provides information on each of the exons (Table 1) contained within the transcripts. The exon information contained in Table 1 includes the coordinates of each exon within its respective transcript, thereby allowing one to readily determine the precise boundaries of each of the exons using the transcript coordinates and the transcript sequences as a reference. These exon boundaries define the exon-exon junctions discussed herein. Also provided in Table 1 is evidence supporting the existence of each exon or transcript (e.g. EST hit, mouse hit, etc.).

Given the transcript sequences provided by the present invention and the exon coordinate information provided in Table 1, or fragments thereof, readily implementable compositions of matter, such as detection elements and detection reagent/kits, (e.g. in the form of probes in a nucleic acid array), can be made using methods well known in the art and discussed herein. Such kits and reagents can be used to track the expression and/or splicing of all of the transcripts/genes disclosed herein, the novel members herein provided, or rationally selected subsets thereof, defined by a user.

Nucleic Acid Arrays and Detection Kits and Reagents

Oligonucleotide probes have long been used to detect complementary nucleic acid sequences in a nucleic acid of interest (the "target" nucleic acid) in the form of detection kits and reagents. In some assay formats, the oligonucleotide probe is tethered, i.e., by covalent attachment, to a solid support, and arrays of oligonucleotide probes immobilized on solid supports have been used to detect specific nucleic acid sequences in a target nucleic acid. See, e.g., PCT patent publication Nos. WO 89/10977 and 89/11548. In other formats, the detection reagents are supplied in solution.

The development of arraying technologies such as photolithographic synthesis of a nucleic acid array and high density spotting of cDNA products has provided methods for making very large arrays of oligonucleotide probes in very small areas. See U.S. Pat. No. 5,143,854 and PCT patent publication Nos. WO 90/15070 and 92/10092. Microfabricated arrays of large numbers of oligonucleotide probes, called "DNA chips" offer great promise for a wide variety of applications. Such arrays may contain, for example, thousands or millions of probes. Probes may

be formed from, for example, cDNA clones, PCR products, or oligonucleotides and can be used in solution or tethered to a support such as a planar surface (chip) or bead format.

The present invention provides detection kits and reagents, such as nucleic acid arrays, that are based on the novel transcript/exon sequences of the human genome provided herein, particularly the novel transcripts and novel information concerning exon structure of each transcript provided in the Sequence Listing and in Table 1.

Medical Importance of Variable Gene Expression

Variable gene expression, such as alternative splicing (also referred to by such terms as alternate splicing or differential splicing) and alternative start/termination sites, is a fundamentally important mechanism of gene regulation. Alternative splicing refers to the formation of two or more different mature mRNA splice forms from a single gene or pre-mRNA, depending on the combination of exons that are spliced together. Alternative splicing therefore serves as an important means of generating additional protein diversity from the structural information encoded by genes. Furthermore, expression of particular splice forms may differ between, for example, cells, tissues, developmental stages/ages, populations or sexes, and may be altered in certain disease states, such as cancer. Alternative splicing may have a detrimental effect on intercellular interactions and the interaction of various polypeptides and cytokines and thereby lead to diseases such as cancer.

Detection reagents, such as nucleic acid arrays and other multi transcript detection reagent/kit, that utilize detection elements comprised of individual transcripts or exons are capable of detecting alternative splice forms of genes that may be missed by detection reagents that detect only one transcript form. Detection reagents that detect disease-specific splice forms of a gene are useful for disease diagnosis. For example, one or more detection reagents to each exon can be used to determine if an exon is present in a sample and/or detection reagents that span exon/exon boundaries can be used to see if a particular exon/exon splice junction is present and also selects against cross reactivity with genomic DNA.

Alternative splicing plays an important role in a variety of proteins and disease pathways, as the following examples illustrate. Elastin is a protein that is important for providing the elastic properties of the lungs, large blood vessels, and skin. The primary elastin transcript undergoes substantial alternative splicing, and it has been suggested that such alternative splicing of elastin may be population-specific and contribute to aging and pathological conditions in the cardiovascular and pulmonary systems (Indik *et al., Am J Med Genet* 1989 Sep;34(1):81-90).

one of the genes/exons of the present invention and novel to the present disclosure. In some kits, such as arrays, the oligonucleotide probes are provided immobilized to a substrate. For example, the same substrate can comprise oligonucleotide probes for detecting at least 1; 10; 100; 1000; 10,000 or most or substantially all of the genes/transcripts or exons provided by the present invention. Any number of probes, or other detection elements, may be utilized in a detection reagent, depending on the particular technology platform and objective. For example, a typical array may contain hundreds, thousands to millions of individual synthetic DNA probes arranged in a grid-like pattern and miniaturized to the size of a dime, each corresponding to a particular exon or transcript/gene. Preferably, probes are attached to a solid support in an ordered, addressable array. Customized arrays that utilize the exon and/or gene/transcript sequences provided by the present invention can be produced by various manufacturers. For example, arrays with over 250,000 oligonucleotide probes or 10,000 cDNAs per square centimeter are readily available (see Lipshutz *et al., Nature Genetics*, 21, 20-24 (1999) and Bowtell *et al., Nature Genetics*, 21, 25-32 (1999). In some arrays, electric fields can be applied to the array to speed hybridization reactions (see Edman *et al., Nucleic Acids Res.* 25, 4907-4914 (1997) and Sosnowski *et al., Proc. Natl. Acad. Sci. USA* 94, 1119-1123 (1997)). Arrays have been previously produced for completely sequenced organisms, such as *Saccharomyces cerevisiae*, that comprise probes for every identified gene in the organism's genome (see DeRisi *et al., Science* 278, 680-686 (1997) and Wodicka *et al., Nature Biotechnology* 15, 1359-1367 (1997)).

The microarray or detection kit is preferably composed of a large number of unique nucleic acid sequences, usually either synthetic antisense oligonucleotides or fragments of cDNAs, fixed to a solid support. Probes may comprise either single- or double-stranded nucleic acid molecules. Oligonucleotides may be about 6-60 nucleotides in length, more preferably 15-30 nucleotides in length, and most preferably about 20-25 nucleotides in length. For a certain type of microarray or detection kit, it may be preferable to use oligonucleotides that are only 7-20 nucleotides in length. For others, such as cDNA, longer lengths are possible and preferable. These can be of the order of 1kb-5kb or more in length and can comprise the entire length of a transcript or exon sequence provided herein or can comprise a short fragment of the transcript/exon, such as in exon-exon junction spanning detection elements.

The microarray or detection kit may contain oligonucleotides that cover, for example, sequential oligonucleotides that cover the full-length sequence, or unique oligonucleotides selected from particular areas along the length of the sequence, such as in exon-exon boundaries. Additionally, such as in the case of primers for PCR, it may be desirable for oligonucleotides to bind to regions 5' or 3' of the transcripts/exons provided herein, such as to capture the entire